

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-307277

(43)Date of publication of application : 22.11.1996

(51)Int.Cl. H03M 7/30  
G10L 9/18  
G11B 20/10

(21)Application number : 07-113361

(71)Applicant : HITACHI LTD

(22)Date of filing : 11.05.1995

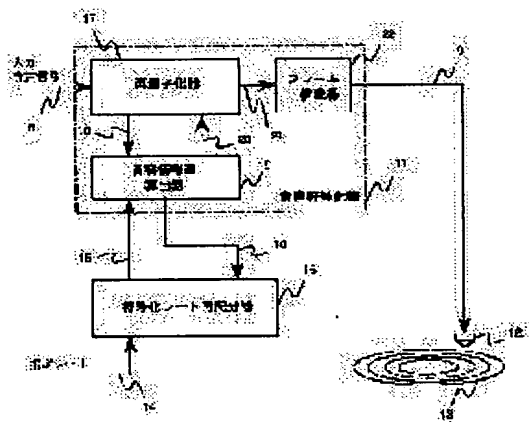
(72)Inventor : OBATA SHINICHI  
NAKAMURA MASAFUMI  
TAKEUCHI TOSHIFUMI

## (54) METHOD AND DEVICE FOR VARIABLE RATE VOICE CODING

## (57)Abstract:

**PURPOSE:** To generate coding data in which best sound quality is reproduced at reproduction by selecting a method of optimum data distribution through the fluctuation of local coding rate based on an information quantity on an audible sense psychology in the voice coder.

**CONSTITUTION:** An acoustic information quantity calculation device 19 calculates acoustic information of a frame acoustic characteristic signal 18 and decides a bit allocation signal 20 in the case of quantization according to the acoustic information quantity. A coding rate re-distributor 16 corrects a distribution of a coding rate in N-frames so that a mean value of the coding rate for the N frames and a designated rate signal 14 are equal to each other and provides an output of a new coding rate 15 for each frame. Bit allocation processing is conducted in the acoustic information quantity calculation device 19 according to the new coding rate to decide the allocation signal 20. The re-quantizer applies re-quantization to a digital audio signal 8 according to the allocation signal 20 and provides an output of a re-quantization signal 21.



## LEGAL STATUS

[Date of request for examination]

06.09.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's]

Best Available Copy

decision of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japanese Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-307277

(43) 公開日 平成8年(1996)11月22日

(51) Int.Cl. <sup>5</sup>	識別記号	片内整理番号	F I	技術表示箇所
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	A
G 1 0 L 9/18			G 1 0 L 9/18	A
G 1 1 B 20/10	3 0 1	7736-5D	G 1 1 B 20/10	3 0 1 Z

審査請求 未請求 請求項の数12 O L (全 13 頁)

(21) 出願番号 特願平7-113361

(22) 出願日 平成7年(1995)5月11日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 小畑 信一

神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所映像メディア研究所内

(72) 発明者 中村 雅文

神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所映像メディア研究所内

(72) 発明者 竹内 敏文

神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所映像メディア研究所内

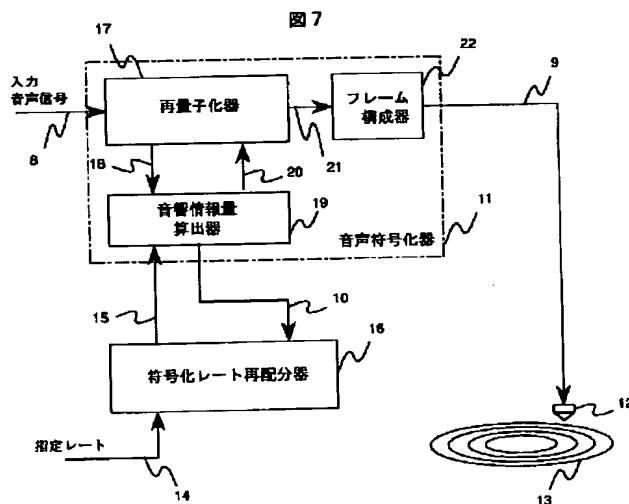
(74) 代理人 弁理士 武 顯次郎

(54) 【発明の名称】 可変レート音声符号化方法及び装置

(57) 【要約】

【目的】 音声符号化装置において、聴覚心理上の情報量をもとに局所符号化レートを変動させて、最適なデータ配分を行う方法を選択することにより、再生時に最良の音質が再現できる符号化データを生成することができる。

【構成】 音響情報量算出器19は、フレーム音響特性信号18の音響情報量を算出し、上記音響情報量に従って量子化に際してのビット割り当て20を決定する。符号化レート再配分器16は、符号化レートのNフレーム間の平均値と指定レート信号14とが等しくなるようにNフレーム内で符号化レートの配分を修正し、フレーム毎の新たな符号化レート15を出力する。この新たな符号化レートに従って音響情報量算出器19内でのビット割り当て処理を行い、割り当て信号20が決定される。再量子化器では上記割り当て信号20に従い、デジタル音声信号8を再量子化し、再量子化信号21を出力する。



## 【特許請求の範囲】

【請求項 1】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もって音響情報量を算出し、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力し、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成し、上記再量子化信号から符号化ビットストリームを構成する音声圧縮符号化方法において、上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くする様に調整することを特徴とする音声符号化方法。

【請求項 2】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もって音響情報量を算出し、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力し、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成し、上記再量子化信号から符号化ビットストリームを構成する音声圧縮符号化方法において、上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くして、且つ予め決めた  $P$  個 ( $P$  は 2 以上の自然数) の符号化単位で構成される音声グループ内で、上記符号化レートの平均値が予め定めた指定符号化レートの値になるように上記符号化レートを制御することを特徴とする音声符号化方法。

【請求項 3】 請求項 1 または 2 記載の音声符号化方法において、音響情報量の算出は、入力デジタル音声信号を周波数軸上で  $N$  個の帯域 ( $N$  は 3 以上の自然数) に分け、帯域  $i$  内 ( $i$  は最低域 0 から最高域  $N-1$  まで) の信号パワー  $S(i)$  と聴覚心理分析をした結果のマスクパワー  $M(i)$  との差  $SMR(i)$  をある決まった帯域  $j$  ( $j$  は 0 以上  $N-2$  以下) からある決まった帯域  $k$  ( $k$  は 1 以上  $N-1$  以下) までに渡って加算したものを上記符号化単位ごとの上記音響情報量として得ることを特徴とする音声符号化方法。

【請求項 4】 請求項 1 または 2 記載の音声符号化方法において、音響情報量の算出は、入力デジタル音声信号を周波数軸上で  $N$  個の帯域 ( $N$  は 3 以上の自然数) に分け、帯域  $i$  内 ( $i$  は最低域 0 から最高域  $N-1$  まで) の信号パワー  $S(i)$  と聴覚心理分析をした結果のマスクパワー  $M(i)$  との差  $SMR(i)$  を得て、更に  $K(i)$  が  $K(i+1)$  以上となるような実数  $K(i)$  を上記  $SMR(i)$  に乗じたものである  $K(i) \times SMR(i)$  をある決まった帯域  $j$  ( $j$  は 0 以上  $N-2$  以下) からある決

まった帯域  $k$  ( $k$  は 1 以上  $N-1$  以下) までに渡って加算したものを上記符号化単位ごとの上記音響情報量として得ることを特徴とする音声符号化方法。

【請求項 5】 請求項 1 または 2 記載の音声符号化方法において、

音響情報量の算出は、入力デジタル音声信号を周波数軸上で  $N$  個の帯域 ( $N$  は 3 以上の自然数) に分け、帯域  $i$  内 ( $i$  は最低域 0 から最高域  $N-1$  まで) の信号パワー  $S(i)$  をある決まった帯域  $j$  ( $j$  は 0 以上  $N-2$  以下) からある決まった帯域  $k$  ( $k$  は 1 以上  $N-1$  以下) までに渡って加算したものを上記符号化単位ごとの上記音響情報量として得ることを特徴とする音声符号化方法。

【請求項 6】 請求項 1 または 2 記載の音声符号化方法において、

音響情報量の算出は、入力デジタル音声信号を周波数軸上で  $N$  個の帯域 ( $N$  は 3 以上の自然数) に分け、帯域  $i$  内 ( $i$  は最低域 0 から最高域  $N-1$  まで) の信号パワー  $S(i)$  に、 $K(i)$  が  $K(i+1)$  以上となるような実数  $K(i)$  を乗じたものである  $K(i) \times S(i)$  をある決まった帯域  $j$  ( $j$  は 0 以上  $N-2$  以下) からある決まった帯域  $k$  ( $k$  は 1 以上  $N-1$  以下) までに渡って加算したものを上記符号化単位ごとの上記音響情報量として得ることを特徴とする音声符号化方法。

【請求項 7】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もって音響情報量を算出し、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力し、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成し、上記再量子化信号から符号化ビットストリームを構成する音声圧縮符号化方法において、上記音響情報量に対してある一定のマージン量  $MARG$  ( $MARG$  は実数) だけ余計にデータ容量が確保できるように上記符号化単位に対する上記符号化レートを決定し、上記符号化単位に対する上記符号化レートの現時点までの平均値  $AVE$  を求め、上記平均値  $AVE$  が予め定めた低レート限界値  $LOW$  より低くなった場合には上記一定のマージン量  $MARG$  を増やし、上記平均値  $AVE$  が予め定めた高レート限界値  $HIGH$  より高くなった場合には上記一定のマージン量  $MARG$  を減らす様にすることを特徴とする音声符号化方法。

【請求項 8】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もり、音響情報量として出力する音響情報量算出手段と、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力する割り当て手段と、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成する再量子化手段と、上記再量子化信号から符号化ビットストリームを構成するストリーム構成手段とを備

## 3

えた音声圧縮符号化装置において、  
上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くする様に調整する符号化レート配分手段を備えることを特徴とする音声符号化装置。

【請求項 9】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もり、音響情報量として出力する音響情報量算出手段と、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力する割り当て手段と、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成する再量子化手段と、上記再量子化信号から符号化ビットストリームを構成するストリーム構成手段とを備えた音声圧縮符号化装置において、  
上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くする様に、且つ予め決めた P 個（P は 2 以上の自然数）の符号化単位で構成される音声グループ内で、上記符号化レートの平均が予め定めた指定符号化レートになるように上記符号化レートを調整する符号化レート配分手段を備えることを特徴とする音声符号化装置。

【請求項 10】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もり、音響情報量として出力する音響情報量算出手段と、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力する割り当て手段と、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成する再量子化手段と、上記再量子化信号から符号化ビットストリームを構成するストリーム構成手段とを備えた音声圧縮符号化装置において、  
上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くする様に、且つ予め決めた P 個（P は 2 以上の自然数）の符号化単位で構成される音声グループ内で、上記符号化レートの平均が予め定めた指定符号化レートになるように調整された符号化レートを設定する符号化レート配分手段と、上記符号化ビットストリームを調整された符号化レートに従った符号化ビットストリームに構成し直す符号化レート変換手段とを備えることを特徴とする音声符号化装置。

【請求項 11】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もり、音響情報量として出力する音響情報量算出手段と、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す

## 4

割り当て信号を出力する割り当て手段と、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成する再量子化手段と、上記再量子化信号から符号化ビットストリームを構成するストリーム構成手段とを備えた音声圧縮符号化装置において、

予め決めた P 個（P は 2 以上の自然数）の符号化単位分の上記デジタル音声信号を溜めるデータ保持手段と、上記符号化単位ごとの上記音響情報量が増大したときには上記符号化単位に対する符号化レートを高くし、減少した場合は上記符号化単位に対する符号化レートを低くする様に、且つ予め決めた P 個（P は 2 以上の自然数）の符号化単位で構成される音声グループ内で、上記符号化レートの平均が予め定めた指定符号化レートになるように調整された符号化レートを設定する符号化レート配分手段とを備え、

上記再量子化手段は上記符号化レート配分手段から出力される上記調整された符号化レートとタイミングを合わせて、上記デジタル音声信号を上記データ保持手段から取り出す様に動作することを特徴とする音声符号化装置。

【請求項 12】 デジタル音声信号を決まった符号化単位ずつ入力し、上記デジタル音声信号の聴覚心理上の情報量を見積もり、音響情報量として出力する音響情報量算出手段と、上記音響情報量に基づいて上記デジタル音声信号を圧縮する際のデータ容量の割当てを示す割り当て信号を出力する割り当て手段と、上記割り当て信号に従って上記デジタル音声信号から再量子化信号を生成する再量子化手段と、上記再量子化信号から符号化ビットストリームを構成するストリーム構成手段とを備えた音声圧縮符号化装置において、

上記音響情報量に対してある一定のマージン量 MARG（MARG は実数）だけ余計にデータ容量が確保できるように上記符号化単位に対する上記符号化レートを決定する符号化レート決定手段と、上記符号化単位に対する上記符号化レートの現時点までの平均値（AVE）を求める符号化レート平均値算出手段と、上記平均値（AVE）が予め定めた低レート限界値（LOW）より低くなった場合には上記一定のマージン量（MARG）を増やし、上記平均値（AVE）が予め定めた高レート限界値（HIGH）より高くなった場合には上記一定のマージン量（MARG）を減らすようにするマージン調整手段とを備えることを特徴とする音声符号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 高密度記録のディスクに音声を記録する音声符号化装置に係り、特に決まった伝送レートや記録媒体上の記録密度の範囲内において、最良の音質を再現できる符号化データを生成する音声圧縮符号化の方法及び装置に関する。

【0002】

## 5

【従来の技術】音質を一定にして高能率の符号化を行う方法の一例としては、特開平4-192724に記載されているものがある。

## 【0003】

【発明が解決しようとする課題】上記の従来方法は、その処理方法の特殊性ゆえにスピーチ音声には有効だが、一般の音響信号、特にCD並を目標にしたデジタルオーディオシステムではその効果をあまり発揮しない。また、現在ある音声圧縮を利用したデジタルオーディオシステムは伝送レートを固定にしているので、アタック音部分では使用可能ビット数の不足により劣化が著しくなる。更に、音声情報量が少ない部分では、必要以上のデータ容量が確保されてしまっている。従って、データ容量の有効活用がされておらず、音質の改善される余地が残っている。

## 【0004】

【課題を解決するための手段】本発明では、入力された音声信号の聴覚心理上の情報量を算出し、その情報量の大小により局所符号化レートを決定する。そしてその局所符号化レートは、予め決めたフレーム単位で平均を取った場合に一定となるように制御する。

## 【0005】

【作用】入力されたデジタル音声信号の聴覚心理上の情報量である音響情報量を音響情報量算出手段が見積もる。この音響情報量が符号化単位ごとに増減するのに合わせて符号化レート配分手段が各符号化単位に対する符号化レートを増減させる。その結果として、音響情報量が多い符号化単位には高い符号化レートが割り当てられ、音響情報量が少ない符号化単位には低い符号化レートが割り当てられる。

## 【0006】

【実施例】図1は本発明の一実施例であり、7つの処理ステップによって構成されている符号化処理フローチャートである。順に処理を説明する。

【0007】処理ステップ1で符号化処理の基本単位であるフレーム毎にデジタル音声信号の入力処理が行われる。次に処理ステップ2で、そのデジタル音声信号に対して、一回目の符号化処理を行うことにより、上記フレーム内のデジタル音声信号に含まれる聴覚心理上の情報量である音響情報量を算出する。

【0008】ここで0から31番までの32個の周波数帯域に分割して音響情報量を算出する例を図2に示す。図2(a)は、信号パワーレベル $S(i)$ と、帯域内マスクや帯域間散布等で聞こえなくなる音の閾値レベルであるマスクレベル $M(i)$ のレベルを各帯域ごとに示したもので、 $S(i)$ と $M(i)$ との差 $SMR(i)$ の大きさが網かけの領域で示してある。図2(b)が $SMR(i)$ を各帯域毎に抜き出したもので、これを帯域 $s_{bmax}$ まで( $s_{bmax}$ は31でなくとも良い)に渡って加算したものを音響情報量として採用している。

## 6

【0009】処理ステップ3では上記一回目の符号化処理による符号化データと上記音響情報量を示すインデックス番号を保存する。処理ステップ4では、 $N$ ( $N$ は2以上の自然数)フレーム分のデータが溜ったかどうかを判断し、溜っていない場合はステップ1に進み、次のフレームの入力処理を行う。溜った場合はステップ5に進む。処理ステップ5では保存されていた $N$ フレーム分のデータ呼び戻す。ステップ6では呼び戻された上記音響情報量を示すインデックス番号により $N$ フレーム内での音響情報量の増減の変化を判断する。そして、個々のフレームの音響情報量に合わせて符号化レートも変換し、且つ $N$ フレーム間の平均符号化レートが予め決まった指定レートになるように、個々のフレームに対する符号化レートを決定する。処理ステップ7で上記決定された符号化レートに従い、本エンコード処理である2回目の符号化処理を行い、データストリームを構成する。

【0010】図3に時間軸に相当するフレーム番号と個々のフレームに対する符号化レートが音響情報量の増減に伴って調整されている例を具体的に示す。図3(a)は図2のような手順に従って求められた、個々のフレームに対する音響情報量の変化の様子を示している。これに対して符号化レートも、図3(b)のように同様の上がり下がり特性を持たせている。これにより、音質は一定となるが、通常、記録媒体の都合等の制約により、一定の記録レートが要求される。そこで、この例では10フレームを1音声グループとして( $N=10$ )、1音声グループ内での平均符号化レートが指定レート値になるように個々のフレームに対する符号化レートが調整されている。

【0011】このように $N$ フレームを単位として、各フレームごとの音響情報量の変化に合わせて符号化レートを決めることで、音質の劣化を平均化し、且つ使用可能なデータ容量の範囲内で最良の音質を保つことができる符号化データの生成が可能となる。

【0012】図4は別の算出法により音響情報量を求めた例であり、各帯域の信号パワー $S(i)$ を帯域 $s_{bmax}$ ( $s_{bmax}$ は31でなくとも良い)までに渡って加算したものを音響情報量として採用している。この例では最小可聴限や帯域間マスクを考慮せずに、単純に帯域毎の音圧レベルの大きさで見積もることになるので高域の信号や小さめの信号も間引かれずに音響情報量に反映する。

【0013】図5は図2のようにして得られた帯域別の $SMR(i)$ に対して帯域ごとに異なる重み付けを施した例である。ここでは、 $K(i)$ が $K(i+1)$ 以上になるような実数 $K(i)$ を用いて、この $K(i)$ と $SMR(i)$ との積を帯域毎に求め、それを帯域 $s_{bmax}$ ( $s_{bmax}$ は31でなくとも良い)までに渡って加算したものを音響情報量として採用している。

【0014】0から31までの帯域が最高周波数までを

等分割したものであるなら、低域ほど1つの帯域が大きいパーク幅を持ち、聴覚心理上大きな影響を与える。そのような場合に対してこの図5のような例の重み付けを行えば、よりの確かな音響情報量の見積りができる。

【0015】図6は、図5と同様の重み付けを図4の例に対して適用したものであり、 $K(i)$ と $S(i)$ との積を帯域毎に求め、それを帯域 $s_{bmax}$  ( $s_{bmax}$ は31でなくとも良い)までに渡って加算したものを音響情報量として採用している。これによってパーク軸上への変換を行わずに、簡易的に聴覚心理上の軸での見積りを行うことができる。

【0016】図7は本発明による符号化装置の一実施例を示したものである。デジタル音声信号8を入力し、符号化ビットストリーム9と音響情報量を示すインデックス10を出力する音声符号化器11、記録ヘッド12、光ディスク、磁気ディスク等の記録媒体13、音響情報量インデックス10と指定レート信号14を入力し、修正符号化レートインデックス15を出力する符号化レート再配分器16、更に音声符号化器11の中身としてデジタル音声信号8を再量子化する再量子化器17とデジタル音声信号8の聴覚心理上の情報量を算出し、使用するべきデータ容量を見積もる音響情報量算出器19と、再量子化された信号21を入力し符号化ビットストリーム9を出力するフレーム構成器22とで構成されている。順に動作を説明する。

【0017】まず第1回目の符号化処理によって音響情報量が見積もられる。符号化の基本単位であるフレーム毎にデジタル音声信号8が音声符号化器11に入力される。音響情報量算出器19は上記1フレームのデジタル音声信号8か或いはデジタル音声信号8を時間周波数変換したスペクトルかであるフレーム音響特性信号18を入力し、聴覚心理上の情報量である音響情報量を算出し、上記音響情報量に従って量子化に際してのビット割り当てを、量子化雑音レベルが検知不能レベル以下となる様に決定する。そしてビット割り当て信号20と音響情報量インデックス10を出力する。

【0018】そして次に同じ音声データの2回目の符号化処理(本エンコード)が行われる。符号化の基本単位であるフレーム毎にデジタル音声信号8が音声符号化器11に入力される。音響情報量算出器19では上記フレーム音響特性信号18の音響情報量を算出し、上記音響情報量に従って量子化に際してのビット割り当て20を決定する。ここで、符号化レート再配分器は1回目の符号化処理で決まった符号化レートのNフレーム間の平均値と指定レート信号14とが等しくなるようにNフレーム内で符号化レートの配分を修正し、フレーム毎の新たな符号化レート15を出力する。この新たな符号化レートに従って音響情報量算出器19内でのビット割り当て処理を行い、割り当て信号20が決定される。そして再量子化器では上記割り当て信号20に従い、デジタ

ル音声信号8を再量子化し、再量子化信号21を出力する。フレーム構成器22は再量子化された信号21から符号化ビットストリーム9を構成して出力する。符号化ビットストリーム9は記録ヘッド12によって記録媒体13に記録される。

【0019】このようにして、1回目の符号化処理でNフレーム内の音響情報量の変化の様子を把握し、2回目の符号化処理で音響情報量に合わせた符号化レートで符号化処理が行われる。この結果、音質の劣化の度合いが平均化されると同時に、制限データ容量内においての最良の音質が得られるビットストリーム9が生成できる。

また、記録媒体13はディスクに限る必要はなく、テープ、その他でも良い。また更に記録媒体に記録するのではなく、伝送路に送出するのでも良い。

【0020】図8は図7での実施例に対し、新たに、入力信号8を溜めておくNフレーム音声メモリ回路23を設けている。

【0021】デジタル音声信号8はメモリ回路23に入力し、音響情報量の見積もりのための音声信号25と本エンコード用の音声信号24とが出力される。音響情報量算出器19は上記音声信号25を入力し、聴覚心理上の情報量である音響情報量を算出し、音響情報量インデックス10を出力する。符号化レート再配分器は音響情報量インデックス10をNフレーム間溜め込み、符号化レートのNフレーム間の平均値と指定レート信号14とが等しくなるように且つ、音響情報量インデックス10の増減に合わせて増減する符号化レート15を決定する。

【0022】Nフレームのそれぞれの符号化レートが決定した後、本エンコード処理として、上記音声信号24が再量子化器17に入力し、上記フレーム音響特性信号18が音響情報量算出器19に送られる。音響情報量算出器19では上記フレーム音響特性信号18の音響情報量を算出し、上記音響情報量に従って量子化に際してのビット割り当て20を決定する。この時、符号化レート15が音声信号24とタイミングが合わされて入力し、符号化レート15の条件の下に割り当て20が決まる。そして再量子化器では上記割り当て信号20に従い、音声信号24を再量子化し、再量子化信号21を出力する。フレーム構成器22は再量子化された信号21から符号化ビットストリーム9を構成して出力する。符号化ビットストリーム9は記録ヘッド12によって記録媒体ディスク13に記録される。

【0023】このようにして、Nフレーム内の音響情報量の変化の様子を把握し、本エンコード処理で音響情報量に合わせて変動する符号化レートで符号化処理が行われる。この結果、音質の劣化の度合いが平均化されると同時に、制限データ容量内においての最良の音質が得られるビットストリーム9が生成できる。

【0024】これによって、音響情報量の見積りと本エ

ンコードを一定時間内で交互に行えるため、短時間でより小さい記録処理遅延時間で図4と同様の符号化処理を行うことができる。

【0025】図9は図7での符号化レート再配分器16の代わりに、マージン量MARG28と音響情報量26を入力し、本来必要とされる音響情報量26に対してマージン量MARG28だけ(MARGは実数)余計にデータ容量が確保できるようにフレームに対する符号化レート29を決定する符号化レート決定器27と、現在までの符号化レート29の平均値AVE31を出力する平均値算出器30と、平均値AVE31と予め決めた上限平均値HIGH32と下限平均値LOW33を入力し、AVE31がHIGH32以上になったらMARG28の値を減らし、AVE31がLOW33以下になったらMARG28の値を増やす様にMARG28の値を制御するマージン調整回路34を備えた例である。

【0026】符号化の基本単位であるフレーム毎にデジタル音声信号8が音声符号化器11に入力される。音響情報量算出器19は上記1フレームのデジタル音声信号8か或いはデジタル音声信号8を時間周波数変換したスペクトルかであるフレーム音響特性信号18を入力し、聴覚心理上の情報量である音響情報量26を算出する。符号化レート決定器27は音響情報量26に対して上記マージン量28だけ余裕が取れるように符号化レート29を決める。符号化レート29に従って音響情報量算出器19は量子化に際してのビット割り当て20を決定する。そしてビット割り当て信号20と音響情報量26を出力する。そして再量子化器では上記割り当て信号20に従い、デジタル音声信号8を再量子化し、再量子化信号21を出力する。フレーム構成器22は再量子化された信号21から符号化ビットストリーム9を構成して出力する。符号化ビットストリーム9は記録ヘッド12によって記録媒体ディスク13に記録される。

【0027】これと並行して、平均値算出器30は符号化レート29の現時点までの平均値AVE31を算出する。マージン調整器34ではAVE31がHIGH32以上になったらMARG28の値を減らし、AVE31がLOW33以下になったらMARG28の値を増やし、MARG28を符号化レート決定器27に送る。

【0028】このようにして、ほぼリアルタイムで可変レート符号化処理ができる。また上記音声グループの切れ目と音響情報量の変化とのタイミングが悪いと、情報量が上限以上の音声グループと下限以下の音声グループが隣接してできてしまう場合があるが、この例では平均値を通算で算出しているため、そのような不都合を回避することができる。また、MARG28は正の数である必要はなく負の数でも良い。MARG28が負の場合はMARG28の絶対値が劣化の度合いを示すが、MARG28が大きいほど、より高い符号化レートが選ばれる点是不会変わらない。

【0029】図10は本発明の他の一例であり、ステップ35、ステップ36、ステップ37の3ステップで構成され、ステップ35は更にステップ38、ステップ39、ステップ40の3ステップで構成され、ステップ37は更にステップ41、ステップ42の2ステップで構成されている符号化処理フローチャートである。順に処理を説明する。

【0030】まず処理ステップ35で1回目の符号化処理が行われる。処理ステップ38で符号化処理の基本単位であるフレーム毎にデジタル音声信号の1回目の入力処理が行われる。次に処理ステップ39で、そのデジタル音声信号に対して、1回目の符号化処理を行うことにより、上記フレーム内のデジタル音声信号に含まれる聴覚心理上の情報量である音響情報量を算出する。

【0031】処理ステップ40では上記1回目の符号化処理による上記音響情報量を示すインデックスを保存する。このステップ38、ステップ39、処理ステップ40がフレーム毎に繰り返される。

【0032】音声信号の1回目の符号化処理が終わったあと、処理ステップ36で保存されていた音響情報量インデックスにより、フレーム毎の音響情報量の増減に合わせて各フレームに対する符号化レートを決定する。

【0033】続いて処理ステップ37で2回目の符号化処理が行われる。処理ステップ41でフレーム毎にデジタル音声信号の2回目の入力処理が行われる。処理ステップ42で上記決定された各フレームに対する符号化レートに従い、本エンコード処理である2回目の符号化処理を行い、データストリームを構成する。このステップ41、ステップ42をフレーム毎に繰り返す。

【0034】このように各フレームごとの音響情報量の変化に合わせて符号化レートを決定することで、音質の劣化を平均化し、且つ最良の音質を保つことができる符号化データの生成が可能となる。

【0035】

【発明の効果】聴覚心理上の情報量に従って符号化レートを決定できるため、ビットがあまり必要のない部分での余裕がビットを多く必要とする部分に回される形となる。これにより、符号化による歪は平均化され且つ最小限に食い止めることができ、最良の音質を保つことができる。

【図面の簡単な説明】

【図1】本発明を適用したデジタル音声符号化方法の処理フローの一実施例である。

【図2】本発明での音響情報量として、SMRの見積りを行う一例である。

【図3】本発明を適用した時の、個々のフレームに対する符号化レートの変化の一例である。

【図4】本発明での音響情報量として、信号パワーの見積りを行う一例である。

【図5】本発明での音響情報量として、帯域毎にSMR



11

に重み付けを加えた物の見積りを行う一例である。

【図 6】本発明での音響情報量として、帯域毎に信号パワーに重み付けを加えた物の見積りを行う一例である。

【図 7】本発明を適用したデジタル音声符号化装置の一実施例である。

【図 8】本発明を適用したデジタル音声符号化装置の一実施例で、入力データメモリ回路を付加した一例である。

【図 9】本発明を適用したデジタル音声符号化装置の一実施例で、リアルタイム処理を可能とした一例である。

12

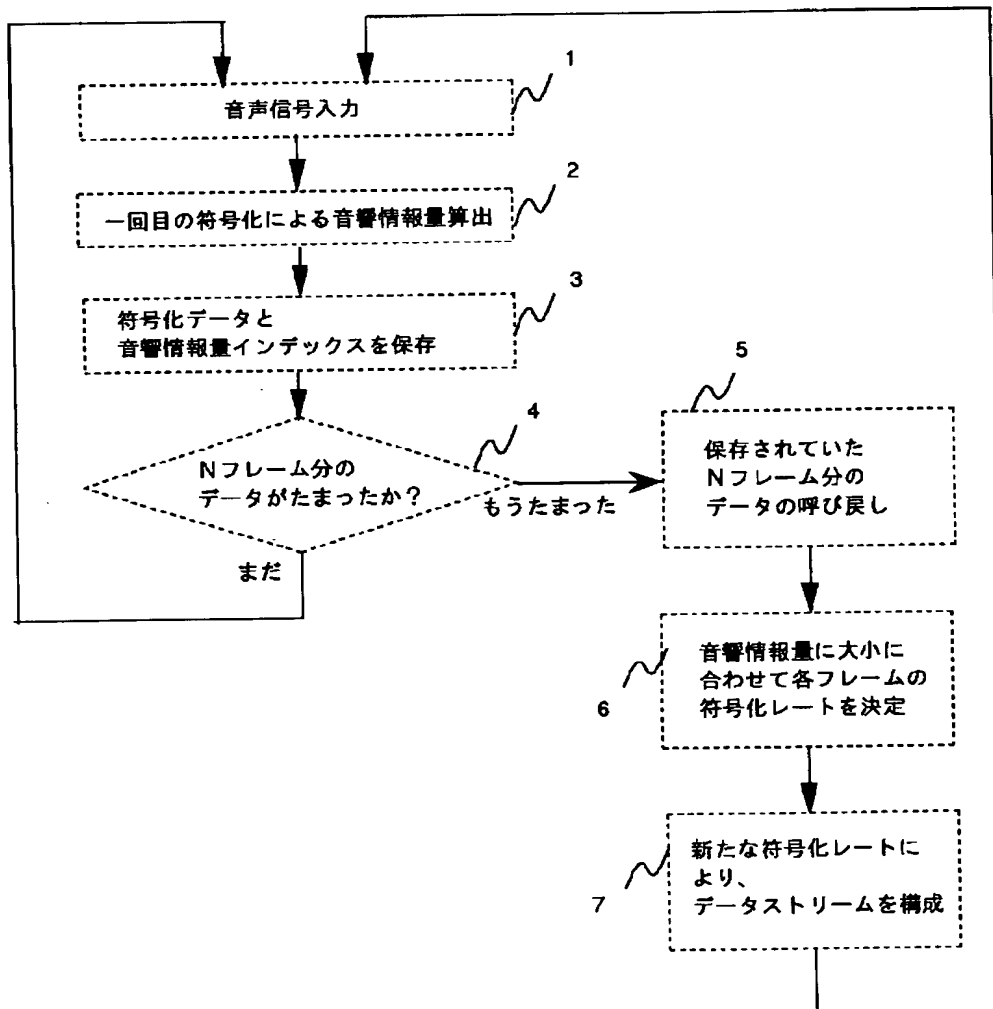
【図 10】本発明を適用したデジタル音声符号化方法の処理フローの他の実施例である。

【符号の説明】

- 8 デジタル音声信号
- 9 符号化ビットストリーム
- 10 音響情報量インデックス
- 11 音声符号化器
- 15 修正符号化レートインデックス
- 16 符号化レート再配分器
- 10 19 音響情報量算出器

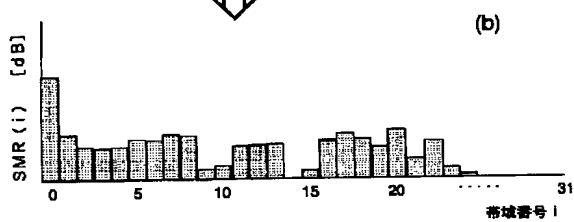
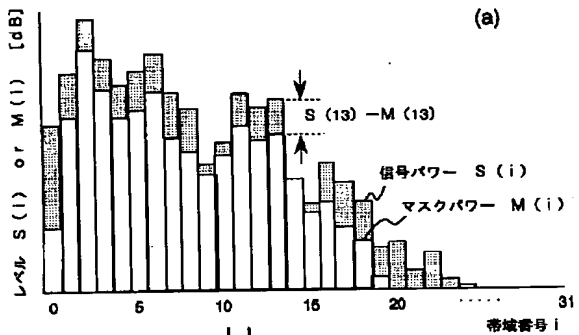
【図 1】

図 1



【図2】

図2

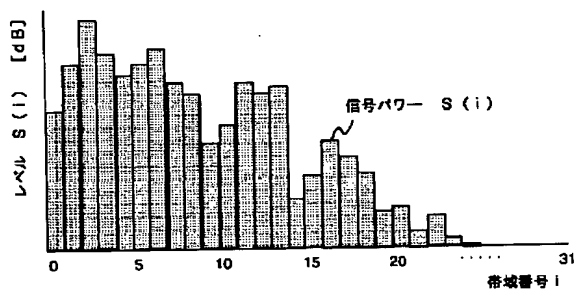


$$\text{音響情報量} = \sum_{i=0}^{sbmax} \{S(i) - M(i)\} = \sum_{i=0}^{sbmax} \text{SMR}(i)$$

但し  $sbmax$  は0以上31以下

【図4】

図4

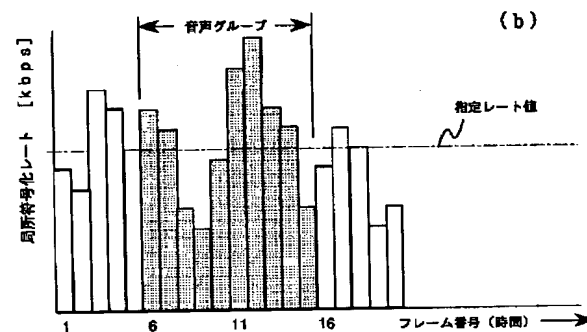
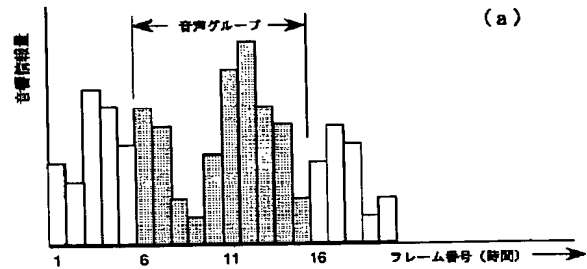


$$\text{音響情報量} = \sum_{i=0}^{sbmax} S(i)$$

但し  $sbmax$  は0以上31以下

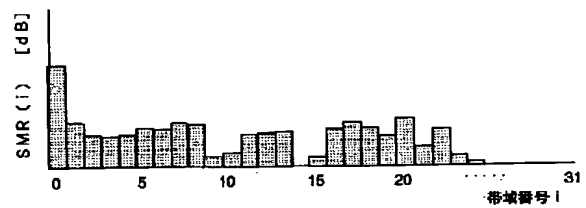
【図3】

図3

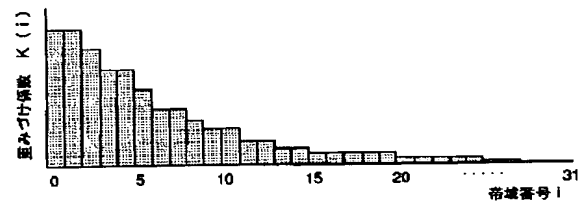


【図5】

図5



X

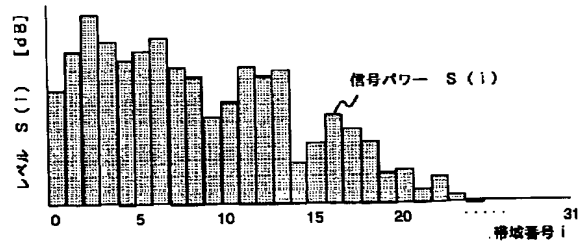


$$\text{音響情報量} = \sum_{i=0}^{sbmax} K(i) \times \text{SMR}(i)$$

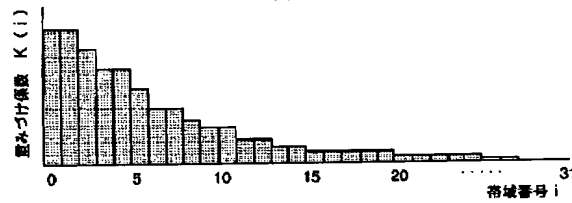
但し  $sbmax$  は0以上31以下

【図6】

図6



×

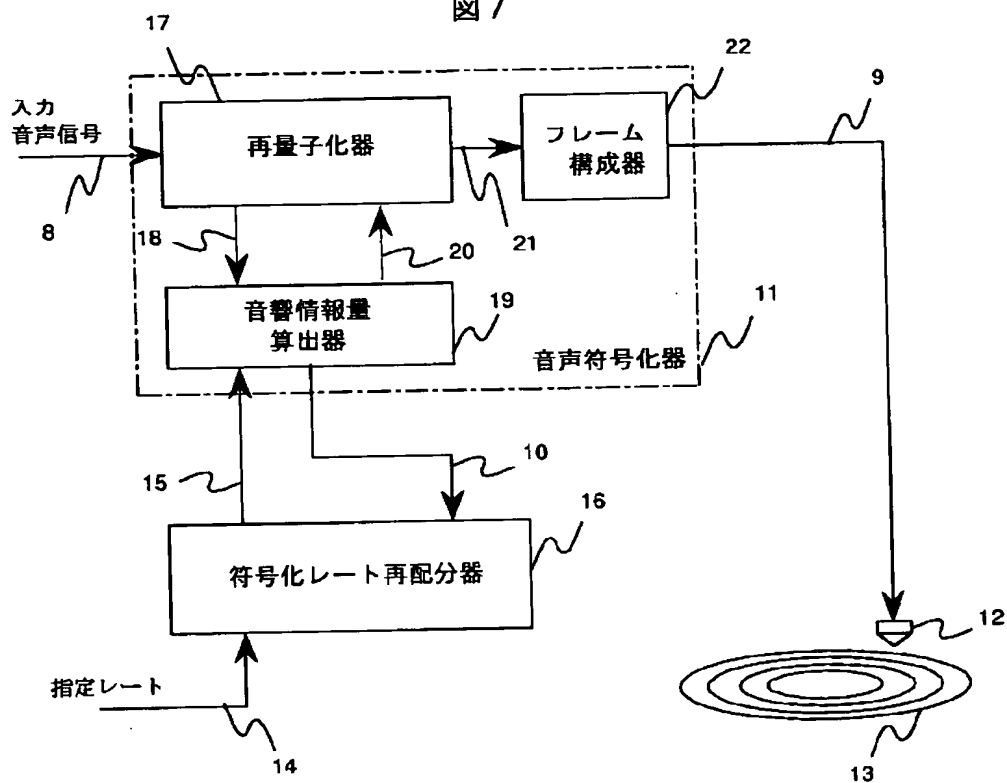


$$\text{音響情報量} = \sum_{i=0}^{sbmax} K(i) \times S(i)$$

但し  $sbmax$  は0以上31以下

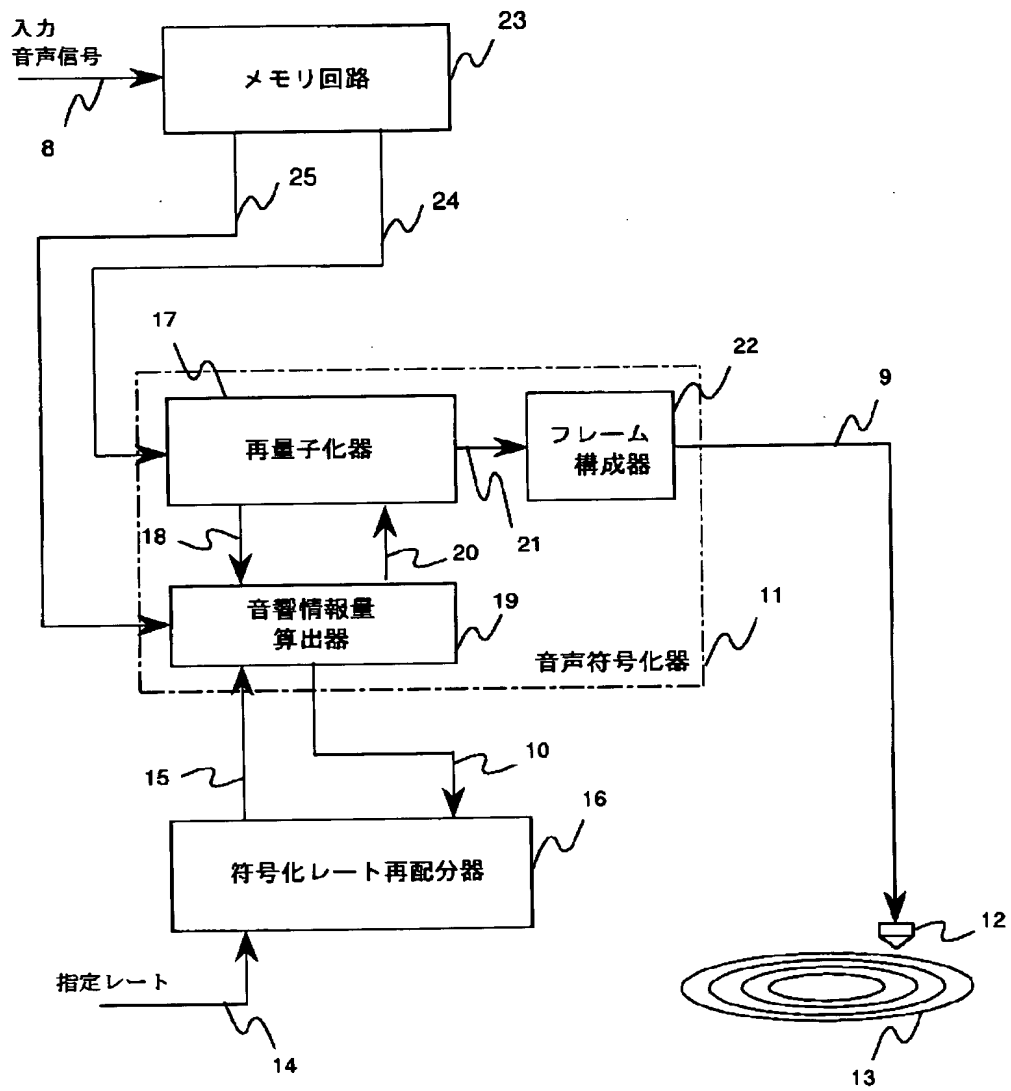
【図7】

図7



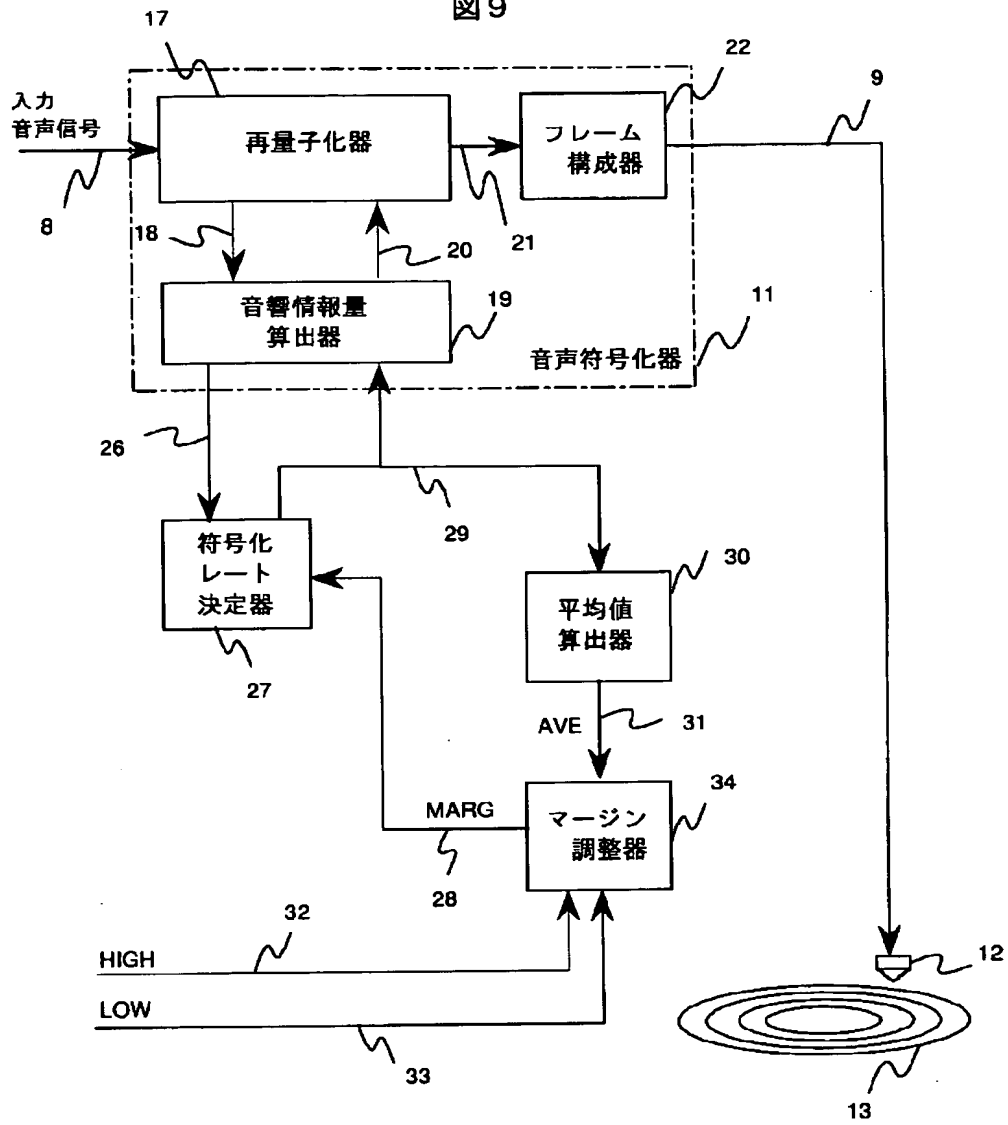
【図 8】

図 8



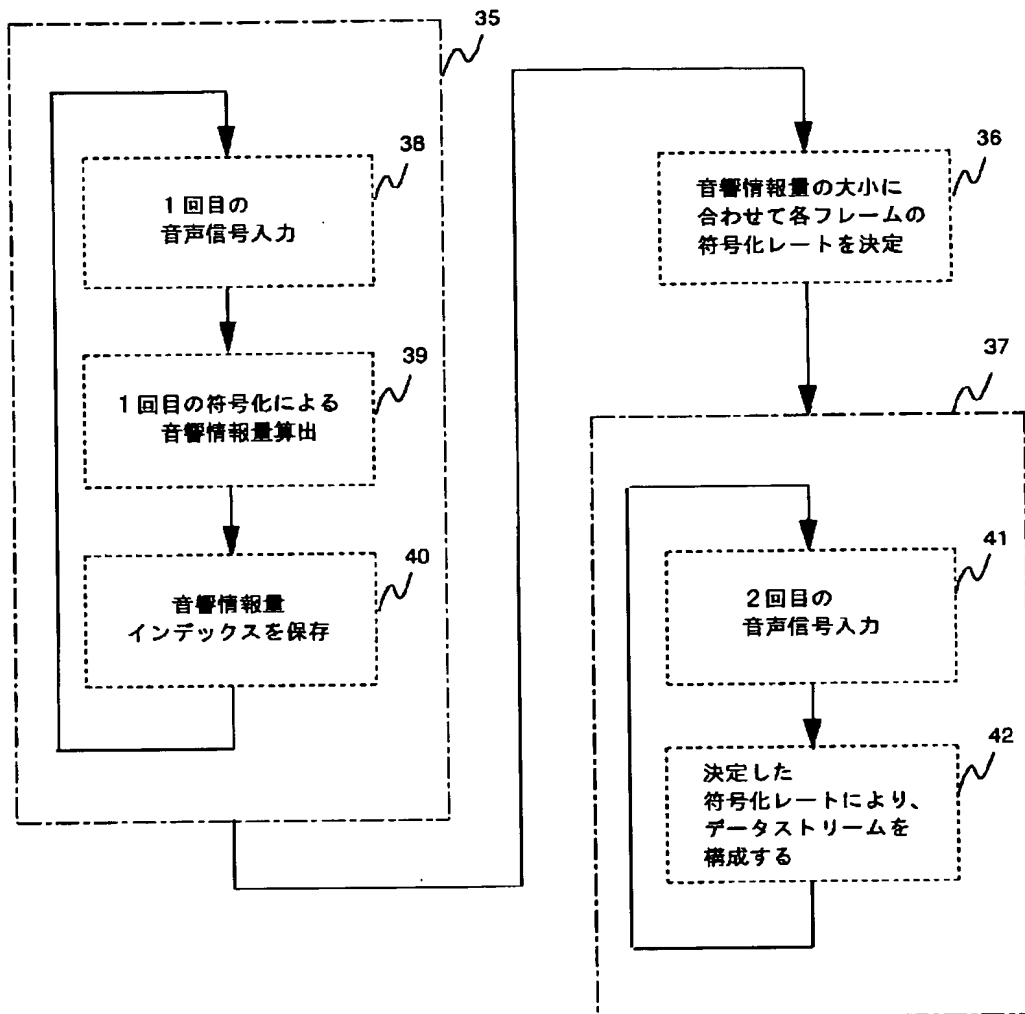
【図 9】

図 9



【図 10】

図 10



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**